

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 10:38:51 ; Search time 15 Seconds

(without alignments)
76.908 Million cell updates/sec

Title: US-09-551-151A-43

Perfect score: 64

Sequence: 1 SPQGIAGQRNFM 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 500 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	65.6	779	1	CGBOIS
2	42	65.6	1042	1	CGCHIS
3	42	65.6	1453	2	S21626
4	42	65.6	1464	1	CGHUIS
5	40	62.5	475	2	T00620
6	40	62.5	1418	2	T43467
7	40	62.5	1419	2	A41182
8	40	62.5	1487	1	CGHUC6
9	40	62.5	1487	2	B41182
10	39	60.9	1320	2	CG1017
11	39	60.9	1320	2	CG1017
12	38.5	60.2	619	2	T64087
13	38	59.4	695	2	T02630
14	38	59.4	702	2	T52634
15	37	57.8	84	2	F50655
16	37	57.8	84	2	G97932
17	37	57.8	442	1	A45390
18	37	57.8	739	2	S47772
19	37	57.8	739	2	D91183
20	37	57.8	739	2	H86029
21	37	57.8	771	2	E98161
22	37	57.8	771	2	AD3126
23	37	57.8	771	2	BA0333
24	37	57.8	1486	1	T24797
25	36.5	57.0	369	2	G83560
26	36	56.2	149	2	S16965
27	36	56.2	258	1	T14704
28	36	56.2	258	1	T14704
29	36	56.2	269	2	T15013

30	36	56.2	324	2	S60118	RING finger proteol
31	36	56.2	383	2	S28422	1,4-alpha-glucan b
32	36	56.2	532	2	B75005	hypothetical prote
33	36	56.2	577	2	AC2790	hypothetical prote
34	36	56.2	598	2	B97569	ABC transporter (A
35	36	56.2	657	2	C71905	probable outer mem
36	36	56.2	660	2	F64608	conserved hypothet
37	36	56.2	677	2	E71879	biotin sulfoxide r
38	36	56.2	826	2	AD0982	1,4-alpha-glucan b
39	36	56.2	830	2	T07824	1,4-alpha-glucan b
40	36	56.2	861	1	S34730	1,4-alpha-glucan b
41	36	56.2	910	2	A34721	androgen receptor
42	36	56.2	911	2	B34721	androgen receptor
43	36	56.2	919	2	A39248	androgen receptor
44	36	56.2	1354	2	T13363	phosphoribosylform
45	36	56.2	1775	2	A31893	collagen alpha 1(I
46	36	56.2	182	2	T27367	hypothetical prote
47	35	54.7	369	2	T24078	hypothetical prote
48	35	54.7	437	2	H86933	aspartate kinase f
49	35	54.7	431	2	AB0129	probable membrane
50	35	54.7	721	2	B83820	hypothetical prote
51	35	54.7	822	2	JT0968	1,4-alpha-glucan b
52	35	54.7	830	2	T06578	1,4-alpha-glucan b
53	35	54.7	914	2	T17233	hypothetical prote
54	35	54.7	924	2	CG3572	organic solvent to
55	35	54.7	1058	2	T30556	aryl hydrocarbon r
56	35	54.7	1059	2	T30557	genome polypeptid
57	35	54.7	1882	1	GNVWTR	collagen alpha 2 f
58	35	54.7	1998	2	A43426	hypothetical prote
59	35	54.7	49	2	A81605	hypothetical prote
60	34	53.1	119	2	B72526	probable membrane
61	34	53.1	125	2	AD0297	hypothetical prote
62	34	53.1	178	2	T29336	hypothetical prote
63	34	53.1	263	2	A83956	flagellar hook pro
64	34	53.1	266	2	B23034	streptomycin resis
65	34	53.1	278	2	T40916	ngl1-interacting f
66	34	53.1	279	2	AC2264	uracil accessory p
67	34	53.1	280	2	AE2026	hypothetical prote
68	34	53.1	283	2	S46770	hypothetical prote
69	34	53.1	286	2	A82929	ATP synthase gamma
70	34	53.1	290	2	A32249	collagen - sea urc
71	34	53.1	370	2	S59539	heat shock transcr
72	34	53.1	393	2	C83049	probable two-compo
73	34	53.1	420	2	D95972	probable sugar upt
74	34	53.1	479	2	F83291	probable outer mem
75	34	53.1	481	2	G86144	hypothetical prote
76	34	53.1	507	2	T34216	exonuclease ABC c
77	34	53.1	579	2	F69157	transcription fact
78	34	53.1	582	2	S10099	probable 3,4-dihyd
79	34	53.1	589	2	F96599	hypothetical prote
80	34	53.1	594	2	T06369	collagen alpha 2(I
81	34	53.1	597	2	B75556	hypothetical prote
82	34	53.1	641	2	S32436	3-isopropylmalate
83	34	53.1	642	2	T30076	transcription fact
84	34	53.1	644	2	JQ0160	collagen alpha 2(I
85	34	53.1	654	2	A34734	hypothetical prote
86	34	53.1	668	2	A53330	transcription fact
87	34	53.1	734	2	T37696	leucyl-cRNA synthet
88	34	53.1	742	4	C34734	leucine-tRNA ligase
89	34	53.1	819	2	G81698	transcription fact
90	34	53.1	825	4	C71544	valine-tRNA ligase
91	34	53.1	829	4	B34734	hypothetical prote
92	34	53.1	881	2	G83811	hypothetical prote
93	34	53.1	910	2	AE3380	valine-tRNA ligase
94	34	53.1	947	2	B97567	valyl-tRNA synthet
95	34	53.1	947	2	A12787	hypothetical prote
96	34	53.1	964	2	T32482	collagen alpha 2 c
97	34	53.1	1051	2	A35763	hypothetical prote
98	34	53.1	1158	2	AF1852	peptide synthetase
99	34	53.1	2459	2	AF2136	hypothetical prote
100	34	53.1	149	2	G96000	hypothetical prote
101	33	51.6	208	2	F70734	hypothetical prote
102	33	51.6				

103	33	51.6	217	2	T45383	176	32	50.0	385	2	G70628	probable acka prot
104	33	51.6	239	2	T07092	177	32	50.0	403	2	B83808	hypothetical prote
105	33	51.6	256	1	RTMST	178	32	50.0	421	2	A71558	probable low calci
106	33	51.6	261	2	T35708	179	32	50.0	440	2	S74953	hypothetical prote
107	33	51.6	264	2	S42534	180	32	50.0	442	1	FOLVUS	gas polypeptide -
108	33	51.6	284	2	T09840	181	32	50.0	448	2	B86182	hypothetical prote
109	33	51.6	293	2	T09920	182	32	50.0	467	2	AC2015	hypothetical prote
110	33	51.6	297	2	T51005	183	32	50.0	478	2	A69477	hypothetical prote
111	33	51.6	303	2	S41754	184	32	50.0	487	2	H87449	IMP dehydrogenase
112	33	51.6	303	2	S58352	185	32	50.0	520	2	S27197	hydroxymethylgluta
113	33	51.6	310	2	I50696	186	32	50.0	520	2	A25332	hydroxymethylgluta
114	33	51.6	321	2	AG2188	187	32	50.0	520	2	S12736	signal recognition
115	33	51.6	324	1	JC4291	188	32	50.0	522	2	A33644	uncharacterized pro
116	33	51.6	324	2	AD2288	189	32	50.0	527	2	C97170	hypothetical prote
117	33	51.6	362	2	D82644	190	32	50.0	533	2	AG2384	very similar to mo
118	33	51.6	385	2	T07130	191	32	50.0	542	2	B90090	nucleoprotein - si
119	33	51.6	397	2	F83132	192	32	50.0	543	2	T01270	hypothetical prote
120	33	51.6	397	2	S13408	193	32	50.0	545	2	T40207	cytochrome-c oxida
121	33	51.6	403	2	T30464	194	32	50.0	551	1	A55582	carboxylesterase (
122	33	51.6	408	2	F69468	195	32	50.0	562	2	A55281	hypothetical prote
123	33	51.6	438	2	S53787	196	32	50.0	572	2	T34273	ABC transporter, A
124	33	51.6	445	2	T20190	197	32	50.0	605	2	T47529	hypothetical prote
125	33	51.6	461	2	F82819	198	32	50.0	631	2	T47529	transcription fact
126	33	51.6	465	2	AD3579	199	32	50.0	638	1	S23391	hypothetical prote
127	33	51.6	470	2	A10883	200	32	50.0	697	2	T16306	collagen COL1 - f
128	33	51.6	481	2	T50691	201	32	50.0	812	2	S31521	Cooc protein precu
129	33	51.6	481	2	F97621	202	32	50.0	872	2	S49582	translation initia
130	33	51.6	481	2	AC2884	203	32	50.0	873	2	T09582	outer membrane ush
131	33	51.6	519	2	E90548	204	32	50.0	901	2	H82850	protein kinase PKM
132	33	51.6	522	2	S13887	205	32	50.0	901	2	S53726	hypothetical prote
133	33	51.6	547	2	H75632	206	32	50.0	902	2	T47966	nuclear pore compl
134	33	51.6	557	2	B83962	207	32	50.0	959	2	B44402	collagen alpha cha
135	33	51.6	572	2	AF0211	208	32	50.0	1027	2	S28774	saccharopine dehyd
136	33	51.6	585	2	E88571	209	32	50.0	1064	2	T05195	endo-1,4-beta-xyla
137	33	51.6	585	2	E88571	210	32	50.0	1077	2	S54975	hypothetical prote
138	33	51.6	636	2	S41067	211	32	50.0	1129	2	T19779	G2R protein - vari
139	33	51.6	820	1	JX0243	212	32	50.0	1264	2	A36858	collagen alpha 1(X
140	33	51.6	886	2	I50694	213	32	50.0	1315	2	A56101	probable SCARECROW
141	33	51.6	966	2	H97717	214	32	50.0	1336	2	T02736	SIR4 protein - yea
142	33	51.6	972	2	S77454	215	32	50.0	1358	2	A29360	collagen alpha 1(Y
143	33	51.6	1049	1	CG8075	216	32	50.0	1492	2	A40333	collagen type V
144	33	51.6	1240	2	T30834	217	32	50.0	1496	1	CGHUV	procollagen type V
145	33	51.6	1414	1	S23809	218	32	50.0	1496	1	S16366	collagen alpha 1(X
146	33	51.6	1464	1	S59856	219	32	50.0	1497	2	I49607	collagen alpha 1(X
147	33	51.6	1466	1	CGH07L	220	32	50.0	1763	2	B56101	collagen alpha 1(X
148	33	51.6	1466	1	I51553	221	32	50.0	1774	2	A34736	collagen alpha 1(X
149	33	50.8	380	2	G85058	222	32	50.0	1805	2	A34736	collagen alpha 1(X
150	32.5	50.8	2471	2	T03820	223	32	50.0	1896	2	B72175	collagen alpha 1(X
151	32	50.0	129	2	B71259	224	32	50.0	1897	2	T28621	collagen alpha 1(X
152	32	50.0	175	2	F72699	225	32	50.0	2022	2	T48818	collagen alpha 1(X
153	32	50.0	183	2	T26637	226	32	50.0	2481	2	A43908	collagen alpha 1(X
154	32	50.0	188	2	A32202	227	32	50.0	2684	2	A96521	collagen alpha 1(X
155	32	50.0	218	2	D64154	228	32	50.0	632	2	B75215	collagen alpha 1(X
156	32	50.0	220	2	D82139	229	32	50.0	632	2	B75215	collagen alpha 1(X
157	32	50.0	221	2	D97161	230	32	50.0	632	2	B75215	collagen alpha 1(X
158	32	50.0	252	2	A47188	231	32	50.0	632	2	B75215	collagen alpha 1(X
159	32	50.0	293	2	F72696	232	32	50.0	632	2	B75215	collagen alpha 1(X
160	32	50.0	297	2	F65002	233	32	50.0	632	2	B75215	collagen alpha 1(X
161	32	50.0	297	2	D91027	234	32	50.0	632	2	B75215	collagen alpha 1(X
162	32	50.0	297	2	E85871	235	32	50.0	632	2	B75215	collagen alpha 1(X
163	32	50.0	304	2	F83642	236	32	50.0	632	2	B75215	collagen alpha 1(X
164	32	50.0	306	2	E97471	237	32	50.0	632	2	B75215	collagen alpha 1(X
165	32	50.0	307	2	T19582	238	32	50.0	632	2	B75215	collagen alpha 1(X
166	32	50.0	308	2	B87059	239	32	50.0	632	2	B75215	collagen alpha 1(X
167	32	50.0	309	2	A95988	240	32	50.0	632	2	B75215	collagen alpha 1(X
168	32	50.0	310	2	T35754	241	32	50.0	632	2	B75215	collagen alpha 1(X
169	32	50.0	311	2	T30905	242	32	50.0	632	2	B75215	collagen alpha 1(X
170	32	50.0	333	2	G83095	243	32	50.0	632	2	B75215	collagen alpha 1(X
171	32	50.0	352	2	C47293	244	32	50.0	632	2	B75215	collagen alpha 1(X
172	32	50.0	361	2	T38638	245	32	50.0	632	2	B75215	collagen alpha 1(X
173	32	50.0	370	2	T42532	246	32	50.0	632	2	B75215	collagen alpha 1(X
174	32	50.0	376	2	T16050	247	32	50.0	632	2	B75215	collagen alpha 1(X
175	32	50.0	380	2	A48295	248	32	50.0	632	2	B75215	collagen alpha 1(X

249	31	48.4	266	2	S61522	LMO protein - frul	322	31	48.4	718	2	G71888	flagellar hook pro
250	31	48.4	267	2	T46087	hypothetical prote	323	31	48.4	718	2	F64628	flagellar hook pro
251	31	48.4	269	2	H65067	hypothetical prote	324	31	48.4	743	2	E84767	hypothetical prote
252	31	48.4	269	2	H91091	probable sensory t	325	31	48.4	743	2	B70017	probable oxidoredu
253	31	48.4	269	2	C85937	probable sensory t	326	31	48.4	749	2	AG2486	exodeoxyribonuclea
254	31	48.4	270	2	UC7631	K+ channel-interac	327	31	48.4	749	2	T15136	hypothetical prote
255	31	48.4	283	2	C75518	streptomycin 3-kin	328	31	48.4	751	4	T43874	hypothetical prote
256	31	48.4	295	2	B41320	hypothetical prote	329	31	48.4	764	2	548521	cathepsin (EC 1.11.
257	31	48.4	304	2	H70769	hypothetical prote	330	31	48.4	809	2	T06457	AKR1 protein - yea
258	31	48.4	311	2	F87655	ABC transporter, A	331	31	48.4	824	2	E87856	probable sucrose s
259	31	48.4	313	2	S30954	minor tail protein	332	31	48.4	846	2	T20710	protein P10D11.6 (
260	31	48.4	316	2	A12122	permease protein o	333	31	48.4	869	2	C56617	hypothetical prote
261	31	48.4	317	2	C82450	l-phosphotransferin	334	31	48.4	895	2	T00787	clac protein precu
262	31	48.4	319	2	AD3227	hypothetical prote	335	31	48.4	906	2	A87413	probable beta-gala
263	31	48.4	319	2	S20086	MyoD protein - sh	336	31	48.4	957	2	H11246	valyl-tRNA synthet
264	31	48.4	322	2	F72800	minor tail subunit	337	31	48.4	971	2	H11719	probable ribonucle
265	31	48.4	324	2	S28672	occr protein - Agr	338	31	48.4	975	2	T28107	hypothetical prote
266	31	48.4	336	2	T19802	hypothetical prote	339	31	48.4	1026	2	T28968	hypothetical prote
267	31	48.4	339	2	S08981	malate dehydrogena	340	31	48.4	1070	2	S75712	cellulase (EC 3.2.
268	31	48.4	342	2	D87305	hypothetical prote	341	31	48.4	1092	2	T45095	probable uridylosy
269	31	48.4	343	2	G95300	hypothetical prote	342	31	48.4	1151	2	S03722	DNA-directed DNA p
270	31	48.4	349	2	T17130	hypothetical prote	343	31	48.4	1159	1	A44280	inner layer proteol
271	31	48.4	352	2	T20729	hypothetical prote	344	31	48.4	1159	2	B98198	hypothetical prote
272	31	48.4	356	1	XNECHC	histidinol-phospha	345	31	48.4	1159	2	AH3088	icmf (imported) -
273	31	48.4	358	1	G75586	urea/short chain-a	346	31	48.4	1274	2	T10729	transferrin-like p
274	31	48.4	366	2	S11449	collagen short cha	347	31	48.4	1286	2	S28634	adhesin AIDA-I pre
275	31	48.4	368	2	A10462	hypothetical prote	348	31	48.4	1618	2	S21424	nestin - human
276	31	48.4	369	2	A87318	hypothetical prote	349	31	48.4	1644	2	F91286	hypothetical prote
277	31	48.4	381	2	T19402	hypothetical prote	350	31	48.4	1644	2	B86128	hypothetical prote
278	31	48.4	384	2	T21502	hypothetical prote	351	31	48.4	1669	1	CGH048	collagen alpha 1(I
279	31	48.4	390	2	AC1893	hypothetical prote	352	31	48.4	1669	1	CGMS48	collagen alpha 1(I
280	31	48.4	410	2	E75190	probable phosphono	353	31	48.4	1781	1	A34374	DNA-directed RNA p
281	31	48.4	411	2	T03294	glutamate dehydrog	354	31	48.4	1797	2	A55677	laminin beta-2 cha
282	31	48.4	411	2	T04342	probable phosphono	355	31	48.4	1798	2	A55677	laminin beta-2 cha
283	31	48.4	412	2	B71222	probable phosphono	356	31	48.4	1804	2	T34518	laminin beta-2 cha
284	31	48.4	413	2	A55238	transcription fact	357	31	48.4	1892	2	C97804	nestin - golden ha
285	31	48.4	416	2	D70347	cell division prot	358	31	48.4	2095	2	S29529	hypothetical prote
286	31	48.4	423	2	A41207	collagen 13, nonfi	359	31	48.4	2524	2	A35844	genome polypeptid
287	31	48.4	424	2	C83902	hypothetical prote	360	31	48.4	2810	2	T22298	Xotch protein - Af
288	31	48.4	427	2	T17123	hypothetical prote	361	31	48.4	3034	2	T14119	hypothetical prote
289	31	48.4	430	2	C89771	Immunoglobulin G b	362	31	48.4	3034	2	S17537	seven-pass transme
290	31	48.4	450	2	B70932	probable PPE prote	363	31	48.4	3034	2	B89939	fibrinolytic prote
291	31	48.4	468	2	T04965	amino acid transpo	364	31	48.4	3034	2	WMBE71	DnaJ protein (Impo
292	31	48.4	473	2	G71345	conserved hypother	365	31	48.4	3034	2	S61185	hypothetical prote
293	31	48.4	482	1	GOBEN4	alkaline exonuclea	366	31	48.4	3034	2	B96914	carbone-monoxide d
294	31	48.4	483	1	T42952	hypothetical prote	367	31	48.4	3034	2	T45238	probable transfe
295	31	48.4	493	2	C87362	hypothetical prote	368	31	48.4	3034	2	AF0472	probable outer mem
296	31	48.4	493	2	T48219	hypothetical prote	369	31	48.4	3034	2	T13590	distal tail fiber
297	31	48.4	502	2	H81203	I-aspartate oxidas	370	31	48.4	3034	2	A27224	myosin heavy chain
298	31	48.4	524	1	OVSNA	hypothetical prote	371	31	48.4	3034	2	B75193	hypothetical prote
299	31	48.4	525	2	T21527	protein A precurs	372	31	48.4	3034	2	A54211	H+-transporting tw
300	31	48.4	528	2	AC3236	hypothetical prote	373	31	48.4	3034	2	JC4912	ectoxin precursor
301	31	48.4	534	2	A99316	hypothetical prote	374	31	48.4	3034	2	T34642	photosystem II pro
302	31	48.4	534	2	A70836	hypothetical prote	375	31	48.4	3034	2	C31183	hypothetical prote
303	31	48.4	545	2	D90159	hypothetical prote	376	31	48.4	3034	2	D84043	probable lipoprote
304	31	48.4	547	2	T39478	zinc-finger protei	377	31	48.4	3034	2	AG0196	hypothetical prote
305	31	48.4	557	2	AD0249	probable hemolysin	378	31	48.4	3034	2	T04286	hypothetical prote
306	31	48.4	561	2	S62788	carboxylesterase (379	31	48.4	3034	2	F70892	probable JpgV prot
307	31	48.4	572	2	S71597	arginyl-tRNA synth	380	31	48.4	3034	2	B70436	hypothetical prote
308	31	48.4	572	2	F81866	arginine-tRNA liga	381	31	48.4	3034	2	T07772	hypothetical prote
309	31	48.4	575	2	S55415	ATP-binding transp	382	31	48.4	3034	2	A69167	disease resistance
310	31	48.4	586	2	C83262	hypothetical prote	383	31	48.4	3034	2	C95018	hypothetical prote
311	31	48.4	587	2	B86311	protein FIL3.16 (l	384	31	48.4	3034	2	A97629	NCRD family protei
312	31	48.4	591	2	JC5797	FOP1-ATPase (EC 3	385	31	48.4	3034	2	AC2852	tRNA/rRNA methylt
313	31	48.4	591	2	T10640	hypothetical prote	386	31	48.4	3034	2	CEFXA	tRNA/rRNA methylt
314	31	48.4	603	1	GBECGC	beta-glucuronidase	387	31	48.4	3034	2	B35127	C-phycocyanin al
315	31	48.4	616	2	T18231	S-lamnin - human	388	31	48.4	3034	2	S18526	phycocyanin alph
316	31	48.4	616	2	T13335	hypothetical prote	389	31	48.4	3034	2	S73038	hypothetical prote
317	31	48.4	669	2	T13335	probable chemotaxi	390	31	48.4	3034	2	H84988	30S ribosomal prot
318	31	48.4	673	2	C83080	hypothetical prote	391	31	48.4	3034	2	C71199	hypothetical prote
319	31	48.4	688	2	E96551	hypothetical prote	392	31	48.4	3034	2	B87593	hypothetical prote
320	31	48.4	693	2	T49296	hypothetical prote	393	31	48.4	3034	2	D83401	hypothetical prote
321	31	48.4	693	2	T49296	hypothetical prote	394	31	48.4	3034	2	S73223	ribosomal protein

395	30	46.9	181	2	A55944	svd protein - Each
396	30	46.9	181	2	E91085	Svd protein limpor
397	30	46.9	181	2	G85930	interacts with sec
398	30	46.9	184	2	S73036	hypothetical prote
399	30	46.9	185	2	S04792	moblization prote
400	30	46.9	187	2	S24992	photosystem II pro
401	30	46.9	200	2	D97891	conserved hypotet
402	30	46.9	201	2	D66912	conserved hypotet
403	30	46.9	201	2	T16181	hypothetical prote
404	30	46.9	202	2	T10016	hypothetical prote
405	30	46.9	202	2	T21165	hypothetical prote
406	30	46.9	216	2	C75403	hypothetical prote
407	30	46.9	216	2	T29039	hypothetical prote
408	30	46.9	218	2	E47188	MHC class II histo
409	30	46.9	221	2	AC0990	probable membrane
410	30	46.9	241	2	F95355	protein limporred
411	30	46.9	251	2	AD2564	hypothetical prote
412	30	46.9	251	2	E71357	probable phosphogl
413	30	46.9	255	1	CPSMMU	mutamoylpentapepti
414	30	46.9	255	1	G83543	conserved hypotet
415	30	46.9	255	2	F83409	hypothetical prote
416	30	46.9	264	2	AD3121	conserved hypotet
417	30	46.9	266	2	C87497	hypothetical prote
418	30	46.9	266	2	T34172	hypothetical prote
419	30	46.9	269	2	A70719	probable enoyl-coA
420	30	46.9	270	2	C98166	hypothetical prote
421	30	46.9	278	2	T16494	hypothetical prote
422	30	46.9	279	2	S62848	H+-transporting tw
423	30	46.9	279	2	C64244	H+-transporting tw
424	30	46.9	281	2	AH1327	nicotinate-nucleot
425	30	46.9	283	2	T29980	hypothetical prote
426	30	46.9	284	2	AC0524	patoclate,beta-alan
427	30	46.9	285	2	A71553	hypothetical prote
428	30	46.9	287	2	T15779	hypothetical prote
429	30	46.9	295	2	T19220	hypothetical prote
430	30	46.9	297	2	T18637	hypothetical prote
431	30	46.9	298	2	T12371	hypothetical prote
432	30	46.9	298	2	E90507	conserved hypotet
433	30	46.9	298	2	S72880	hypothetical prote
434	30	46.9	300	2	F83621	hypothetical prote
435	30	46.9	308	2	A70761	hypothetical prote
436	30	46.9	311	2	AG2409	mRNA-binding prote
437	30	46.9	315	2	H83096	probable pyrophosp
438	30	46.9	316	2	F97644	UDP-hexose transfe
439	30	46.9	316	2	AB2868	UDP-hexose transfe
440	30	46.9	318	2	AG3576	iron(III) dictiraf
441	30	46.9	318	2	AH2189	hypothetical prote
442	30	46.9	325	2	T18594	hypothetical prote
443	30	46.9	328	2	T36574	probable partition
444	30	46.9	345	1	C69966	hypothetical prote
445	30	46.9	345	1	S34503	photosystem II pro
446	30	46.9	348	2	AB3260	hypothetical membr
447	30	46.9	352	1	HCHU	alpha-1-microglobu
448	30	46.9	352	2	A22780	photosystem II pro
449	30	46.9	352	2	S14137	photosystem II pro
450	30	46.9	353	1	FMS332	photosystem II pro
451	30	46.9	353	1	FMMT3D	photosystem II pro
452	30	46.9	353	1	FMMH32	photosystem II pro
453	30	46.9	353	1	FMSX3N	photosystem II pro
454	30	46.9	353	1	FMMT32	photosystem II pro
455	30	46.9	353	1	FMS732	photosystem II pro
456	30	46.9	353	1	FMR332	photosystem II pro
457	30	46.9	353	1	F2VFD1	photosystem II pro
458	30	46.9	353	1	F2BWD1	photosystem II pro
459	30	46.9	353	1	F2BWD1	photosystem II pro
460	30	46.9	353	1	F2NND1	photosystem II pro
461	30	46.9	353	1	F2CNU1	photosystem II pro
462	30	46.9	353	1	F2PMD1	photosystem II pro
463	30	46.9	353	1	F2NT1C	photosystem II pro
464	30	46.9	353	1	F2KMTM	photosystem II pro
465	30	46.9	353	1	F2LVD1	photosystem II pro
466	30	46.9	353	1	F2MMD1	photosystem II pro
467	30	46.9	353	2	T07195	photosystem II pro

468	30	46.9	353	2	S29327	photosystem II pro
469	30	46.9	353	2	S15957	photosystem II pro
470	30	46.9	353	2	A21730	photosystem II pro
471	30	46.9	353	2	A25579	photosystem II pro
472	30	46.9	353	2	A25580	photosystem II pro
473	30	46.9	353	2	S33912	photosystem II pro
474	30	46.9	353	2	S57265	photosystem II pro
475	30	46.9	353	2	S58531	photosystem II pro
476	30	46.9	353	2	S44245	photosystem II pro
477	30	46.9	353	2	S42492	photosystem II pro
478	30	46.9	355	2	G82941	hypothetical prote
479	30	46.9	356	2	F90881	histidinol-phospha
480	30	46.9	356	2	D85827	probable lipoprote
481	30	46.9	358	2	F70577	photosystem II pro
482	30	46.9	360	1	F2KTD1	photosystem II pro
483	30	46.9	360	1	F2A117	photosystem II pro
484	30	46.9	360	1	F2A112	photosystem II pro
485	30	46.9	360	1	F2YB16	photosystem II pro
486	30	46.9	360	1	F2YB17	photosystem II pro
487	30	46.9	360	1	F2YB13	photosystem II (b
488	30	46.9	360	2	T11963	photosystem II pro
489	30	46.9	360	2	S73133	photosystem II pro
490	30	46.9	360	2	A25362	photosystem II pro
491	30	46.9	360	2	S26586	photosystem II pro
492	30	46.9	360	2	A20978	photosystem II pro
493	30	46.9	360	2	S25362	photosystem II pro
494	30	46.9	360	2	S32577	photosystem II pro
495	30	46.9	360	2	S32576	photosystem II pro
496	30	46.9	360	2	S45009	photosystem II pro
497	30	46.9	360	2	S54256	photosystem II pro
498	30	46.9	360	2	A48306	photosystem II pro
499	30	46.9	360	2	S78284	photosystem II pro
500	30	46.9	360	2	AB2414	photosystem II pro

ALIGNMENTS

RESULT 1

CGBOIS collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence: revision 31-Dec-1993 #text: change 31-Mar-2000

C:Accession: A91193; A91229; A91387; A91211; A91201; A91200; A43048; A02853

R:Rauterberg, J.; Timpi, R.; Furtmayr, H.

Eur. J. Biochem. 27, 231-237, 1972

A>Title: Structural characterization of N-terminal antigenic determinants in calf and

A:Reference number: A91193; MUID:72255334; PMID:4115172

A:Accession: A91193

A:Molecule type: protein

A:Residues: 1-19 <RND>

A:Experimental source: skin

A>Note: The epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is co

R:Rauterberg, J.; Timpi, R.; Furtmayr, H.

Eur. J. Biochem. 52, 77-82, 1975

A>Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-brom

A:Reference number: A91229; MUID:7602320; PMID:1164916

A:Accession: A91229

A:Molecule type: protein

A:Residues: 20-145 <FIE>

A:Experimental source: skin

A>Note: Lys-103 is hydroxylated and binds glucosylgalactose

R:Rauterberg, J.; Timpi, R.; Furtmayr, H.

FEBS Lett. 26, 74-76, 1972

A>Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from c

A:Reference number: A91211; MUID:74086118; PMID:4359390
 A:Accession: A91211
 A:Molecule type: protein
 A:Residues: 295-562 <F13>
 A:Experimental source: skin
 R:Wendt, P.; Mark, K.V.D.; Rexrodt, F.; Kuehn, K.
 Eur. J. Biochem. 30, 169-183, 1972
 A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues
 A:Reference number: A91201; MUID:73042276; PMID:4343808
 A:Accession: A91201
 A:Molecule type: protein
 A:Residues: 563-675 <MEN>
 A:Experimental source: skin
 R:Fietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.
 Eur. J. Biochem. 30, 163-168, 1972
 A:Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB6-C
 A:Reference number: A91200; MUID:73042275; PMID:4343807
 A:Accession: A91200
 A:Molecule type: protein
 A:Residues: 676-758 <F14>
 A:Experimental source: skin
 A:Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in position
 R:Rautenberg, J.; Fietzek, P.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.
 FEBS Lett. 21, 75-79, 1972
 A:Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of
 A:Reference number: A43048
 A:Accession: A43048
 A:Molecule type: protein
 A:Residues: 759-779 <RA2>
 A:Experimental source: skin
 C:Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydroxylated
 C:Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) are H
 9, 149, 268, and 217 residues.
 C:Comment: The complete chain contains 1052 residues.
 C:Superfamily: collagen alpha 1(I) chain: fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer;
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 65.6%; Score 42; DB 1; Length 779;
 Best Local Similarity 100.0%; Pred. No. 7.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 POGIAGOR 9
 |||||
 Db 513 POGIAGOR 520

RESULT 2
 C:Species: Gallus gallus (chicken) (tentative sequence) (fragments)
 C:Date: 12-Aug-1981 #sequence revision 06-Jul-1982 #text change 31-Mar-2000
 C:Accession: A90458; A90181; A02857
 R:Highberger, J.H.; Corbett, C.; Dixit, S.N.; Yu, W.; Seyer, J.M.; Kang, A.H.; Gross, J.
 Biochemistry 21, 2048-2055, 1982
 A:Title: Amino acid sequence of chick skin collagen alpha1(I)-C88 and the complete prime
 A:Reference number: A90458; MUID:82231995; PMID:7093229
 A:Accession: A90458
 A:Molecule type: protein
 A:Residues: 1-1036 <HIG>
 A:Experimental source: skin
 A:Note: This is the latest in a series of papers from these workers elucidating the sequ
 R:Eyre, D.R.; Glimmer, M.J.
 Biochem. Biophys. Res. Commun. 48, 720-726, 1972
 A:Title: Evidence for a previously undetected sequence at the carboxyterminus of the alp
 A:Reference number: A90181; MUID:72243016; PMID:5047697
 A:Accession: A90181
 A:Molecule type: protein
 A:Residues: 1037-1042 <EVR>
 A:Experimental source: skin
 A:Note: residues 1037-1042 above correspond to the carboxyl end of the protein
 C:Comment: Lysines at positions 103, 700, 934, and 946 above may be hydroxylated in some

C:Comment: Most of the prolines at the third position of the tripeptide repeating unit
 C:Comment: Pro-1002 is the only 3-hydroxyproline and the only hydroxylated proline in
 C:Superfamily: collagen alpha 1(I) chain: fibrillar collagen carboxyl-terminal homolo
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trime
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 65.6%; Score 42; DB 1; Length 1042;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 POGIAGOR 9
 |||||
 Db 789 POGIAGOR 796

RESULT 3
 S21626
 collagen alpha 1(I) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence revision 25-Apr-1997 #text change 13-Aug-1999
 C:Accession: S57243; S16374; A23982; I49559; I49557; S39789; I48300; S21626
 R:Li, S.W.; Killian, J.; Prockop, D.J.
 Matrix Biol. 14, 593-595, 1994
 A:Title: The complete cDNA coding sequence for the mouse pro-alpha-1(I) chain of type
 A:Reference number: S57243
 A:Accession: S57243
 A:Molecule type: mRNA
 A:Residues: 1-1453 <LUS>
 A:Cross-references: EMBL:U08020; NID:q470673; PIDN:AAA88912.1; PID:q470674
 R:Metaeater, M.; Tomen, D.; de Crombrughe, B.; Vuorio, E.
 Biochim. Biophys. Acta 1089, 241-243, 1991
 A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNA
 A:Reference number: S16176; MUID:91274555; PMID:2050384
 A:Accession: S16374
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1442-1453 <MET>
 A:Cross-references: EMBL:X57981; NID:q550484; PIDN:CAA41046.1; PID:q550485
 R:French, B.T.; Lee, W.H.; Maul, G.C.
 Gene 39, 311-312, 1985
 A:Title: Nucleotide sequence of a cDNA clone for mouse proalpha1(I) collagen protein.
 A:Reference number: A23982; MUID:86137403; PMID:3841523
 A:Accession: A23982
 A:Molecule type: mRNA
 A:Residues: 518-1128 <FRE>
 A:Cross-references: GB:M14423; NID:q192261; PIDN:AAA37333.1; PID:q192262
 R:Monson, J.M.; Friedman, J.; McCarthy, B.J.
 Mol. Cell. Biol. 2, 1362-1371, 1982
 A:Title: DNA sequence analysis of a mouse pro-alpha-1(I) procollagen gene: Evidence f
 A:Reference number: I49559; MUID:83141374; PMID:6298597
 A:Accession: I49559
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 735-1130 <RES>
 A:Cross-references: GB:M17491; NID:q192263; PIDN:AAA37334.1; PID:q192264
 R:Hatters, K.; Kuehn, M.; Dellus, H.; Jaenisch, R.
 Proc. Natl. Acad. Sci. U.S.A. 81, 1504-1508, 1984
 A:Title: Insertion of retrovirus into the first intron of alpha1(I) collagen gene lea
 A:Reference number: I49557; MUID:84170331; PMID:6324198
 A:Accession: I49557
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-25 <RE2>
 A:Cross-references: GB:K01688; NID:q192246; PIDN:AAA37330.1; PID:q553881
 R:Fenton, S.P.; Lamande, S.R.; Hannigan, M.; Stacey, A.; Jaenisch, R.; Bateman, J.F.
 Biochim. Biophys. Acta 1216, 469-474, 1993
 A:Title: Genomic sequence of mouse COL1A1 encoding the collagen propeptides.
 A:Reference number: S39789; MUID:94092741; PMID:8268229
 A:Accession: S39789
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-185, 1031-1201, 'G', 1203-1218, 'E', 1220-1221, 'T', 122
 R:Rhodes, K.; Rippe, R.A.; Umezawa, A.; Nehls, M.; Brenner, D.A.; Breindl, M.
 Mol. Cell. Biol. 14, 5950-5960, 1994

A:Title: DNA methylation represses the murine alpha 1(I) collagen promoter by an indirect
 A:Reference number: 148300; MUID:94344105; PMID:8065328
 A:Accession: 148300
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REF>
 A:Cross-references: EMBL:X54676; NID:950486; PIDN:CAA3657.1; PID:950487
 C:Genetics:
 A:Gene: COL1A1
 A:Introns: 770/3; 788/3; 806/3; 842/3; 860/3; 878/3; 932/3; 968/3; 1004/3; 1022/3; 1058/
 A:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; heterotrimer; triple helix
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-151/Domain: amino-terminal propeptide #status predicted <PRO>
 F:30-89/Domain: von Willebrand factor type C repeat homology <VWC>
 F:152-1453/Product: collagen alpha 1(I) chain #status predicted <MAT>
 F:1224-1453/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 65.6%; Score 42; DB 2; Length 1453;
 Best local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIACOR 9
 |||||
 DB 940 POGIACOR 947

RESULT 4
 CGH015
 collagen alpha 1(I) chain precursor - human
 N:Alternate names: procollagen alpha 1(I) chain
 C:Species: Homo sapiens (man)
 C:Date: 12-Aug-1981 #sequence revision 04-oct-1996 #text change 31-Dec-2000
 C:Accession: 160114; S01143; A93335; 155254; A39943; 155237; A55233; S09400; B90567; S11
 5269; A29439; 153466; A02852; 137247
 R:D'Alessio, M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.
 Gene 67, 105-115, 1988
 A:Title: Complete nucleotide sequence of the region encompassing the first twenty-five
 A:Reference number: 160114; MUID:88329734; PMID:2843432
 A:Accession: 160114
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-369, 'U', 371-589 <DAL>
 A:Cross-references: GB:M0789; NID:9179593; PIDN:AA59373.1; PID:9179594
 R:Tromp, G.; Kuitavanti, H.; Stacey, A.; Shikata, H.; Baldwin, C.T.; Jaenisch, R.; Prock
 Blochem. J. 253, 919-922, 1988
 A:Title: Structure of a full-length cDNA clone for the prepro-alpha-1(I) chain of human
 A:Reference number: S01143; MUID:89025644; PMID:3178743
 A:Accession: S01143
 A:Molecule type: mRNA
 A:Residues: 1-472 <TRO>
 A:Cross-references: EMBL:X07884; NID:930015; PIDN:CAA30731.1; PID:930016; GB:M36546; NID
 A:Note: Submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
 R:Chu, M.L.; de Wet, W.; Bernard, M.; Ding, J.F.; Morabito, M.; Myers, J.; Williams, C.;
 Nature 310, 337-340, 1984
 A:Title: Human proalpha1(I) collagen gene structure reveals evolutionary conservation of
 A:Reference number: A93335; MUID:84270697; PMID:6462220
 A:Accession: A93335
 A:Molecule type: DNA
 A:Residues: 1-58, 'Q', 60-181 <CHD>
 A:Cross-references: EMBL:X00820; NID:935657; PIDN:CAA25394.1; PID:935658
 R:Rossoff, C.M.S.; Vergeer, W.P.; du Plooy, S.J.; Bernard, M.P.; Ramirez, F.; de Wet, W.
 J. Biol. Chem. 262, 15151-15157, 1987
 A:Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene en
 A:Reference number: 155254; MUID:88033098; PMID:2822714
 A:Accession: 155254
 A:Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-45 <ROS>
 A:Cross-references: NID:9180387; PIDN:AA51993.1; PID:9180388
 R:Bornstein, P.; McKay, J.; Morishima, J.K.; Devareyalu, S.; Gelinas, R.E.
 Proc. Natl. Acad. Sci. U.S.A. 84, 8869-8873, 1987
 A:Title: Regulatory elements in the first intron contribute to transcriptional control

A:Reference number: A39943; MUID:88097389; PMID:3480516
 A:Accession: A39943
 A:Molecule type: DNA
 A:Residues: 1-34 <BOC>
 A:Cross-references: GB:J03559; NID:9180876; PIDN:AA52052.1; PID:9553338
 R:Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F.
 J. Biol. Chem. 260, 2315-2320, 1985
 A:Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promote
 A:Reference number: 155237; MUID:85130970; PMID:2857713
 A:Accession: 155237
 A:Status: translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-34 <CHD>
 A:Cross-references: GB:M10627; NID:9180383; PIDN:AA51993.1; PID:9553226
 R:Witz, M.K.; Keene, D.R.; Holt, H.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Holl
 J. Biol. Chem. 265, 6312-6317, 1990
 A:Title: In vivo and in vitro noncovalent association of excised alpha1(I) amino-term
 rome, type VII.
 A:Reference number: A35233; MUID:90202908; PMID:2318655
 A:Accession: A35233
 A:Molecule type: protein
 A:Residues: 33-52 <WTR>
 A:Note: This propeptide fragment remained non-covalently bound to a defective, unclea
 R:Well, D.; d'Alessio, M.; Ramirez, F.; de Wet, W.; Cole, W.G.; Chan, D.; Bateman, J.
 EMBO J. 8, 1705-1710, 1989
 A:Title: A base substitution in the exon of a collagen gene causes alternative splici
 A:Reference number: S09400; MUID:89356643; PMID:2767050
 A:Accession: S09400
 A:Molecule type: mRNA
 A:Residues: 156-183 <WEI>
 R:Click, E.M.; Bornstein, P.
 Biochemistry 9, 4699-4706, 1970
 A:Title: Isolation and characterization of the cyanogen bromide peptides from the alp
 A:Reference number: A90567; MUID:71038625; PMID:5529814
 A:Contents: CNBR0-1, CNBR2, CNBR4, CNBR5
 A:Accession: B90567
 A:Molecule type: protein
 A:Residues: 162-188, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233,
 A:Experimental source: skin
 A:Note: evidence for 170-alanine
 R:Baetge, B.; Notbohm, H.; Diebold, J.; Lehmann, H.; Bodo, M.; Deutzmann, R.; Mueller
 Eur. J. Biochem. 192, 153-159, 1990
 A:Title: A critical crosslink region in human bone-derived collagen type I. Specific
 A:Reference number: S11372; MUID:90382436; PMID:2169412
 A:Accession: S11372
 A:Molecule type: protein
 A:Residues: 175-187, 274-287, 'P', 289 <BAE>
 A:Note: Sequence of collagen alpha 1(S)(I) isolated from bone after pepsin digestion
 R:Deak, S.B.; Scholtz, P.M.; Amenta, P.S.; Constantinou, C.D.; Levi-Mizel, S.A.; Gonza
 J. Biol. Chem. 266, 21827-21832, 1991
 A:Title: The substitution of arginine for glycine 85 of the alpha 1(I) procollagen ch
 A:Reference number: 155342; MUID:92042092; PMID:1718984
 A:Accession: 155342
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 258-268; 1347-1357 <DEA>
 A:Cross-references: GB:567495; NID:9239007; PIDN:AA20350.1; PID:9239008
 A:Note: sequences from the 5' and 3' ends only are shown; mutant sequence 263-Arg rep
 R:Morgan, P.H.; Jacobs, H.G.; Segrest, J.P.; Cunningham, L.W.
 J. Biol. Chem. 245, 5042-5048, 1970
 A:Title: Comparative study of glycopeptides derived from selected vertebrate collagen
 A:Reference number: A92069; MUID:771001508; PMID:4319110
 A:Accession: A92069
 A:Molecule type: protein
 A:Residues: 263-268 <MOR>
 A:Experimental source: skin
 A:Note: attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose to 5-hydroxylysine
 R:Radhard, M.E.; Hollister, D.W.
 Matrix 10, 124-130, 1990
 A:Title: Segmental amplification of the entire helical and telopeptide regions of the
 A:Reference number: S15989; MUID:90326017; PMID:2374517
 A:Accession: S15989

```

R:Comn, D.H.; Apone, S.; Eyre, D.R.; Starman, B.J.; Andreassen, P.; Charbonneau, H.;
J. Biol. Chem. 263, 14605-14607, 1988
A:Title: Substitution of Cysteine for Glycine within the Carboxyl-terminal Telopeptide
A:Reference number: I55269; MUID:89008319; PMID:3170557
A:Accession: I55269
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1187-1194, 'C', 1196-1220 <CON>
A:Cross-references: GB:M3213; MID:g340842; PIDN:AAB59363.1; PID:g499622
A>Note: mutant sequence from a patient with mild osteogenesis imperfecta
R:Makekela, J.K.; Raasina, M.; Virta, A.; Vuorio, E.
Nucleic Acids Res. 16, 349, 1988
A:Title: Human pro-alpha-1(I) collagen: cDNA sequence for the C-propeptide domain.

Query Match 65.6%; Score 42; DB 1; Length 1464;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PGIAGQR 9
|||||
Db 951 PGIAGQR 958

RESULT 5
T00620
probable amino acid transport protein T2711.3 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 22-Oct-1999
C:Accession: T00620
R:Feederpiel, N.A.; Palm, C.J.; Conway, A.B.; Kurts, D.B.; Conway, A.R.; Au, M.; Arau
Vytasala, V.; Yu, G.; Ecker, J.; Theologis, A.; Davis, R.W.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z14193
A:Accession: T00620
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-475 <FED>
A:Cross-references: EMBL:AC004122; MID:g3176693; PIDN:AC34329.1; PID:g3540179; GSPDB
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:r2711.3
A:Map position: 1
A:Insertions: 108/1; 139/2; 211/1; 258/3; 334/3
C:Superfamily: Arabidopsis amino acid transport protein I
C:Keywords: amino acid transport; transmembrane protein; transport protein

Query Match 62.5%; Score 40; DB 2; Length 475;
Best Local Similarity 58.3%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SPQIAGORNFN 12
|||||
Db 87 SPDSITGRNRYN 98

RESULT 6
T45467
collagen alpha 1(II) chain precursor [imported] - horse
N:Alternate names: type II collagen
C:Species: Equus caballus (domestic horse)
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 04-Mar-2000
C:Accession: T45467
R:Richardson, D.W.; Dodge, G.R.
submitted to the EMBL Data Library, June 1996
A:Description: Cloning of equine type II collagen and modulation of its expression in
A:Reference number: Z23977
A:Accession: T45467
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1418 <RIC>
A:Cross-references: EMBL:U62528; PIDN:AAB05773.1
A:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

```


Query Match 62.5%; Score 40; DB 2; Length 1418;
 Best Local Similarity 87.5%; Pred. No. 33;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIACOR 9
 111111111
 Db 904 POGIACOR 911

RESULT 7
 A:Title: collagen alpha 1(II) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 13-Aug-1999
 C:Accession: A41182; A44885
 R:Metaepranta, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.
 J. Biol. Chem. 266, 16862-16869, 1991
 A:Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, and
 A:Reference number: A41182; MUID:91358489; PMID:1885613
 A:Accession: A41182
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-1419 <MET>
 A:Cross-references: GB:M65161
 R:Cheah, K.S.; Lau, E.T.; Au, P.K.; Tam, P.P.
 Development 111, 945-953, 1991
 A:Title: Expression of the mouse alpha 1(II) collagen gene is not restricted to cartilag
 A:Reference number: A44885; MUID:91347939; PMID:1879363
 A:Accession: A44885
 A:Molecule type: DNA
 A:Residues: 1-28 <CH2>
 A:Cross-references: GB:S63190; NID:9234368; PIDN:AA19627.1; PID:9234369
 A:Note: sequence extracted from NCBI backbone (NCBIN:63190; NCBI:63192)
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; trime
 F:1191-1419/domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 62.5%; Score 40; DB 2; Length 1419;
 Best Local Similarity 87.5%; Pred. No. 33;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIACOR 9
 111111111
 Db 905 POGIACOR 912

RESULT 8
 CGH06C
 A:Title: collagen alpha 1(II) chain precursor [validated] - human
 N:Alternate names: procollagen alpha 1(II) chain
 N:Context: chondrocalcin; collagen alpha 1(II) chain precursor splice form 1; collagen
 C:Species: Homo sapiens (man)
 C:Date: 28-May-1986 #sequence_revision 01-Sep-1995 #text_change 08-Dec-2000
 C:Accession: A38513; S06715; S24270; A24828; S06496; A35428; A30147; A33116; S64674; S63
 7250; I37251; I37252; I37253; I37254; I35388; I35933; I61910
 R:Ryan, M.C.; Sieraski, M.; Sandell, L.J.
 Genomics 8, 41-48, 1990
 A:Title: The human type II procollagen gene: identification of an additional protein-cod
 A:Reference number: A38513; MUID:91184811; PMID:2081599
 A:Accession: A38513
 A:Molecule type: DNA
 A:Residues: 1-103 <RYA>
 A:Cross-references: GB:M60299; NID:9180883; PIDN:AA173873.1; PID:9180884
 R:Su, M.W.; Lee, B.; Ramirez, F.; Machado, M.; Horton, W.
 Nucleic Acids Res. 17, 9473, 1989
 A:Title: Nucleotide sequence of the full length cDNA encoding for human type II procolla
 A:Reference number: S06715; MUID:90067946; PMID:2587267
 A:Accession: S06715
 A:Molecule type: mRNA
 A:Residues: 1-28, 'R', '99-1487, <SU2>
 A:Cross-references: EMBL:X16468; NID:929515; PIDN:CA34488.1; PID:929516
 A:Note: alternative splice form 1

R:Vikku, M.; Metsaeranta, M.; Syvaenen, A.C.; Ala-Kokko, L.; Vuorio, E.; Peltonen,
 Biochem. J. 285, 287-294, 1992
 A:Title: Structural analysis of the regulatory elements of the type-II procollagen ge
 A:Reference number: S24270; MUID:92344585; PMID:1637314
 A:Accession: S24270
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-28 <VIR>
 A:Cross-references: EMBL:X58709; GB:S40537; NID:935659
 R:Nunez, A.M.; Kono, K.; Martin, G.R.; Yamada, Y.
 A:Note: this translation is not annotated in GenBank entry HSPROCOL, release 111.0
 Gene 44, 11-16, 1986
 A:Title: Promoter region of the human pro-alpha-1(II)-collagen gene.
 A:Reference number: A24828; MUID:87031574; PMID:3021582
 A:Accession: A24828
 A:Molecule type: DNA
 A:Residues: 1-8, 'T', '10-28 <NUN>
 A:Cross-references: GB:M25698; NID:9180872; PIDN:AA52051.1; PID:9553237
 R:Baldwin, C.T.; Reginato, A.M.; Smith, C.; Jimenez, S.A.; Prockop, D.J.
 Biochem. J. 267, 521-528, 1989
 A:Title: Structure of cDNA clones coding for human type II procollagen. The alpha-1(I
 A:Reference number: S06496; MUID:90026318; PMID:2803268
 A:Accession: S06496
 A:Molecule type: mRNA
 A:Residues: 7-28, 'R', '99-157, 'P', '159-440, 'G', '442-456, 'E', '458-640, 'A', '642-831, 'PA', '834,
 A:Cross-references: EMBL:X16711; NID:930040; PIDN:CA34683.1; PID:930041
 A:Note: alternative splice form 1
 R:Ryan, M.C.; Sandell, L.J.
 J. Biol. Chem. 265, 10334-10339, 1990
 A:Title: Differential expression of a cysteine-rich domain in the amino-terminal prop
 A:Reference number: A35428; MUID:90285153; PMID:2355003
 A:Accession: A35428
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 27-81, 'L', '83-103 <RYA2>
 A:Note: alternative splice form 2; splicing appears to be under developmental regulat
 R:Su, M.W.; Benson-Chandy, V.; Vissing, H.; Ramirez, F.
 Genomics 4, 438-441, 1989
 A:Title: Organization of the exons coding for Pro alpha-1(II) collagen N-propeptide c
 A:Reference number: A30147; MUID:89233138; PMID:2714801
 A:Accession: A30147
 A:Molecule type: DNA
 A:Residues: 104-157, 'P', '159-236 <SUM>
 A:Cross-references: GB:J03065; GB:IM23660; GB:M25655; GB:M25656; GB:M25730; GB:M32168;
 R:Ala-Kokko, L.; Baldwin, C.T.; Moskowicz, R.W.; Prockop, D.J.
 Proc. Natl. Acad. Sci. U.S.A. 87, 6565-6568, 1990
 A:Title: Single base mutation in the type II procollagen gene (COL2A1) as a cause of
 A:Reference number: A94227; MUID:90370826; PMID:1975693
 A:Accession: A33116
 A:Molecule type: DNA
 A:Residues: 171-172, 'C', '174-175 <ALA>
 A:Note: mutant sequence from a family with family with primary generalized osteoartr
 R:Diab, M.; Wu, J.J.; Eyre, D.R.
 Biochem. J. 314, 327-332, 1996
 A:Title: Collagen type IX from human cartilage: a structural profile of intermolecula
 A:Reference number: S64673; MUID:96195147; PMID:8660302
 A:Accession: S64674
 A:Molecule type: Protein
 A:Residues: 188-189, 'X', '191-195; 1224-1230, 'X', '1232-1236 <DIA>
 R:Franc, S.; Marzin, E.; Bouillon, M.M.; Latont, R.; Lechene de la Porte, P.; Herbag
 Eur. J. Biochem. 234, 125-131, 1995
 A:Title: Immunohistochemical and biochemical analyses of 20000-25000-year-old fossil
 A:Reference number: S63514; MUID:96096730; PMID:8529631
 A:Accession: S63514
 A:Molecule type: Protein
 A:Residues: 243-261; 575-590; 756-763, 'X', '765-779 <FRA>
 R:Iller, G.E.; Weis, M.A.; Polunbo, P.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Ey
 Am. J. Hum. Genet. 56, 388-395, 1995
 A:Title: An RNA-splicing mutation (G451V520) in the type II collagen gene (COL2A1) in
 A:Reference number: I38867; MUID:95150028; PMID:7847372
 A:Accession: I38867
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 440,'G',442-456,'E',458-480,'P',482-509 <THII>
 A:Cross-references: EMBL:U15195; NID:9557053; PIDN:AA60370.1; PID:9557054
 R:Ramirez, F.
 submitted to the EMBL Data Library, December 1988
 A:Reference number: 504892
 A:Accession: 504892
 A:Molecule type: mRNA
 A:Residues: 501-676,'A',678-783,'A',785-831,'PA',834,'F',836-1214 <RAN>
 A:Cross-references: EMBL:X13783; NID:930037; PIDN:CAA2030.1; PID:930030
 R:Vilkula, M.; Peltonen, L.
 FEBS Lett. 250, 171-174, 1989
 A:Title: Structural analysis of the polymorphic area in type II collagen gene.
 A:Reference number: 505000; MUID:89325561; PMID:2753125
 A:Accession: 505000
 A:Molecule type: DNA
 A:Residues: 630-640,'A',642-765 <YIK2>
 A:Cross-references: EMBL:X16158; NID:929951; PIDN:CAA34278.1; PID:91335018; PIDN:CAA3427
 R:Bogetoft, R.; Tiller, G.E.; Weis, M.A.; Gruber, H.E.; Rimoin, D.L.; Cohn, D.H.; Eyre, D.
 J. Biol. Chem. 267, 22522-22526, 1992
 A:Title: An amino acid substitution (Gly853->261u) in the collagen alpha 1(II) chain pro
 A:Reference number: A44309; MUID:93054548; PMID:1429602
 A:Accession: A44309
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: DNA; mRNA
 A:Residues: 752-831,'PA',834,'F',836-1005,'K',1007-1036,'Q',1038-1052,'E',1054-1068,'T',
 A:Cross-references: GB:100977; NID:9180812; PIDN:AA623914.1; PID:9258774
 A:Note: sequence extracted from NCBI backbone (NCBI:117273); parts of this sequence wer
 A:Note: this translation is not annotated and this publication is not cited in Genbank
 R:Tiller, G.E.; Rimoin, D.L.; Murray, L.W.; Cohn, D.H.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3889-3893, 1990
 A:Title: Tandem duplication within a type II collagen gene (COL2A1) exon in an individual
 A:Reference number: S16502; MUID:90251662; PMID:2339128
 A:Accession: S16502
 A:Molecule type: DNA
 A:Residues: 1164-1184,'GPGSGKDCANCIPEPT',1185-1199 <TL2>
 A:Cross-references: EMBL:M37126; NID:9180808; PIDN:AA52037.1; PID:9180809
 A:Note: mutant sequence from a patient with spondyloepiphyseal dysplasia
 R:Chen, K.S.E.; Stoker, N.G.; Griffin, J.R.; Grosved, F.G.; Solomon, E.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2555-2559, 1985
 A:Title: Identification and characterization of the human type II collagen gene (COL2A1)
 A:Reference number: A02858; MUID:85190534; PMID:3857598
 A:Accession: A02858
 A:Molecule type: DNA
 A:Residues: 1032-1056,'N',1058-1068,'T',1070-1487 <CHE>
 A:Cross-references: GB:J00116; NID:9180395; PIDN:AA5197.1; PID:9180396
 R:Elima, K.; Vuorio, T.; Vuorio, E.
 Nucleic Acids Res. 15, 9489-9504, 1987
 A:Title: Determination of the single polyadenylation site of the human pro-alpha-1(II) c
 A:Reference number: A27280; MUID:88067771; PMID:2825137
 A:Accession: A27280
 A:Molecule type: mRNA
 A:Residues: 1175-1487 <ELI>
 A:Cross-references: EMBL:X06568; NID:930096; PIDN:CAA29504.1; PID:930097
 R:van der Rest, M.; Rosenberg, L.C.; Olsen, B.R.; Poole, A.R.
 Biochem. J. 237, 923-925, 1986
 A:Title: Chondrocalcin is identical with the C-propeptide of type II procollagen.
 A:Reference number: A57033; MUID:87099927; PMID:3800925
 A:Accession: A57033
 A:Molecule type: protein
 A:Residues: 'XE',1244-1246,'N',1248,'X',1250-1265,1295-1305,1395-1408 <VAN>
 A:Note: Chondrocalcin identified as released collagen 1(II) chain carboxyl-terminal pro
 R:Strom, C.M.; Upholt, W.B.
 Nucleic Acids Res. 12, 1025-1038, 1984
 A:Title: Isolation and characterization of genomic clones corresponding to the human ty
 A:Reference number: A21733; MUID:84118798; PMID:6320112
 A:Accession: A21733
 A:Molecule type: DNA
 A:Residues: 1245-1295 <STR1>
 A:Cross-references: EMBL:X00339; EMBL:X00298; NID:934659; PIDN:CAA25092.1; PID:94378975
 A:Accession: B21733

A:Molecule type: DNA
 A:Residues: 894-909,'PE' <STR2>
 A:Cross-references: GB:K01785; NID:930035; PIDN:CAA25082.1; PID:91335032
 R:Nunez, A.M.; Francmanno, C.; Young, M.F.; Martin, G.R.; Yamada, Y.
 Biochemistry 24, 6343-6348, 1985
 A:Title: Isolation and partial characterization of genomic clones coding for a human
 gene.
 A:Reference number: A24561; MUID:86104139; PMID:3002437
 A:Accession: A24561
 A:Molecule type: DNA
 A:Residues: 1296-1358 <NUN2>
 A:Cross-references: GB:M12048; NID:9180017
 A:Note: this translation is not annotated in Genbank entry HUKCCT2A, release 111.0
 A:Note: the codons given for 1333-Tys (AGG) and 1350-Gly (GCA) are inconsistent with
 R:Sanjorgi, F.O.; Benson-Chanda, V.; de Wet, W.J.; Sobel, M.E.; Tsipouras, P.; Ramir
 Nucleic Acids Res. 13, 2207-2225, 1985
 A:Title: Isolation and partial characterization of the entire human pro alpha 1(II) c
 A:Reference number: 137249; MUID:85215609; PMID:2987845
 A:Accession: S59491
 A:Molecule type: DNA
 A:Residues: 7-28,'R',99-114,541-578,786-802,1055-1056,'N',1058-1068,'T',1070-1109,120
 A:Accession: 184453
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 7-28 <SAN2>
 A:Cross-references: GB:M3735; NID:9180845; EMBL:X03320; GB:M24938; NID:930104
 A:Note: the Genbank PID is based on an incorrect reading frame
 A:Accession: 137250
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 541-560 <SAN3>
 A:Cross-references: EMBL:X02378; GB:M23870; NID:930107; PIDN:CAA26227.1; PID:9929621
 A:Accession: 137251

Query Match 62.5% Score 40; DB 1; Length 1487;
 Best Local Similarity 87.5% Pred. No. 35;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 POGIAGOR 9
 Db 973 POGIAGOR 980

RESULT 9
 B41182
 collagen alpha 1(II) chain precursor (long splice form) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 16-Jul-1999
 C:Accession: B41182
 R:Metzgerant, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.
 J. Biol. Chem. 266, 16862-16869, 1991
 A:Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, a
 A:Reference number: A41182; MUID:91358489; PMID:1885613
 A:Accession: B41182
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-1487 <MET>
 A:Cross-references: GB:M65161
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo
 C:Keywords: alternative splicing; coiled coil; extracellular matrix; glycoprotein; tr
 F:33-91/Domain: von Willebrand factor type C repeat homology <WVC>
 F:1259-1487/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 62.5% Score 40; DB 2; Length 1487;
 Best Local Similarity 87.5% Pred. No. 35;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 POGIAGOR 9
 Db 973 POGIAGOR 980

RESULT 10

881961
 Phosphoribosylformylglycinamide synthase (EC 6.3.5.3) NMA0445 [imported] - *Neisseria meningitidis*
 C:Species: *Neisseria meningitidis*
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
 C:Accession: G81961
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holtz, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A:Reference number: AB1775; MUID:2022356; PMID:10761919
 A:Accession: G81961
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1320 <PAR>
 A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83743.1; PID:g737919
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: putL; NMA0445
 C:Superfamily: phosphoribosylformylglycinamide synthase
 C:Keywords: ligase

Query Match 60.9%; Score 39; DB 2; Length 1320;
 Best Local Similarity 80.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SPOGIAGORN 10
 Db 1265 SPOGIAGVTN 1274

RESULT 11
 G81017
 Phosphoribosylformylglycinamide synthase NMB1996 [imported] - *Neisseria meningitidis*
 C:Species: *Neisseria meningitidis*
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: G81017
 R:Reitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Halt, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Rlin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, B.; Kous, H.; Niemann, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
 A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A:Reference number: AB1000; MUID:20175755; PMID:10710307
 A:Accession: G81017
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1320 <TEXT>
 A:Cross-references: GB:AE002549; GB:AE002098; NID:g7227258; PIDN:AAF42323.1; PID:g722725
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1996
 C:Superfamily: phosphoribosylformylglycinamide synthase

Query Match 60.9%; Score 39; DB 2; Length 1320;
 Best Local Similarity 80.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 SPOGIAGORN 10
 Db 1265 SPOGIAGVTN 1274

RESULT 12
 I64087
 translation elongation factor EF-seib homolog - *Haemophilus influenzae* (strain Rd KW20)
 C:Species: *Haemophilus influenzae*
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
 C:Accession: I64087; T09413
 R:Reichmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glöck, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A:Reference number: A64000; MUID:95350630; PMID:7542800
 A:Accession: I64087
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-619 <TIGR>
 A:Cross-references: GB:U32753; GB:L42023; NID:g1573701; PIDN:AMC2366.1; PID:g1573710
 C:Genetics:
 A:Gene: HI0709
 C:Superfamily: translation elongation factor seib; translation elongation factor Tu h
 C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis; selenocyst
 E:1-114/Domain: translation elongation factor Tu homology <ETU>
 E:7-14/Region: nucleotide-binding motif A (P-loop)

Query Match 60.2%; Score 38.5; DB 2; Length 619;
 Best Local Similarity 66.7%; Pred. No. 27;
 Matches 10; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

OY 1 SPOGIAGOR---NEN 12
 Db 223 SPOGIAGORLNLN 237

RESULT 13
 T02630
 hypothetical protein At2g25930 [imported] - *Arabidopsis thaliana*
 N:Alternate names: hypothetical protein T19L18.26
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
 C:Accession: T02630; E84654
 R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke, submitted to the EMBL Data Library, August 1998
 A:Description: *Arabidopsis thaliana* chromosome II BAC T19L18 genomic sequence.
 A:Reference number: Z14681
 A:Accession: T02630
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-695 <KOU>
 A:Cross-references: EMBL:AC004747; NID:g3413096; PID:g3413719
 A:Experimental source: cultivar Columbia
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, B.; Kous, H.; Niemann, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: AB4420; MUID:20083487; PMID:10617197
 A:Accession: E84654
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-695 <STO>
 A:Cross-references: GB:AE002093; NID:g3413719; PIDN:AMC31242.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g25930; T19L18.26
 A:Map position: 2
 A:Introns: 73/3; 339/2; 356/3

Query Match 59.4%; Score 38; DB 2; Length 695;
 Best Local Similarity 60.0%; Pred. No. 38;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 POGIAGORNF 11
 Db 608 POGISGSKSF 617

RESULT 14
 T52634
 nematode responsive protein [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
 C:Accession: T52634
 R:Puzio, P.S.; Tauben, J.; Almeida-Engler, J.; Cai, D.; Gheysen, G.; Grondler, G.M.W.; Gene 239, 163-172, 1999

A:Title: Isolation of a gene from Arabidopsis thaliana related to nematode feeding stru
A:Reference number: 226146
A:Accession: J52634
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-702 <PUZ>
A:Cross-references: EMBL:Y11994; PIDN:CAA72719.1
A:Experimental source: cell line C24; shoot; adult plant
C:Function:
A:Description: functions probably as transcription factor

Query Match 59.4%; Score 38; DB 2; Length 702;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 POGIACORNF 11
||||:|:|
DB 608 POGISGSKSF 617

RESULT 15

F95065
hypothetical protein SP0563 [Imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: F95065
R:Retelien, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hel-
son, J.D.; Hickey, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Kadane, D.; Holtzapple,
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21557209; PMID:11463916
A:Accession: F95065
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 184 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74719.1; PID:g14972038; GSPDB:GN00164; TIGR:SPA
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0563

Query Match 57.8%; Score 37; DB 2; Length 84;
Best Local Similarity 50.0%; Pred. No. 6;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 POGIACORNF 11
||:|:|:|
DB 75 PEGIFGERNY 84

RESULT 16

G97932
hypothetical protein spr0487 [Imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: G97932
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E-
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M-
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: G97932
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK99291.1; PID:g15458058; GSPDB:GN00174
C:Genetics:
A:Gene: spr0487

Query Match 57.8%; Score 37; DB 2; Length 84;

Best Local Similarity 50.0%; Pred. No. 6;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 POGIACORNF 11
||:|:|:|
DB 75 PEGIFGERNY 84

RESULT 17

A45390
gag polyprotein - Maedi/Visna virus (strain KV1772) (provirus)
N:Alternate names: core polyprotein
N:Contains: core protein p14; core protein p16; core protein p25
C:Species: Maedi/Visna virus
A:Note: host Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: A45390
R:Andersson, O.S.; Elser, J.E.; Tobin, G.J.; Greenwood, J.D.; Gonda, M.A.; Georgsson,
J.W.; Petrusson, G.
Virology 193, 89-105, 1993
A:Title: Nucleotide sequence and biological properties of a pathogenic proviral molec
A:Reference number: A45390; MUID:93174981; PMID:8382414
A:Accession: A45390
A:Molecule type: DNA
A:Residues: 1-442 <AND>
A:Cross-references: GB:S55323; NID:9265825; PIDN:AA25459.1; PID:9265826
C:Genetics:
A:Gene: gag
A:Superfamily: AIDS-related virus gag polyprotein
C:Keywords: core protein; polyprotein
F:1-143/Product: core protein p16 #status predicted <P16>
F:144-363/Product: core protein p25 #status predicted <P25>
F:364-442/Product: core protein p14 #status predicted <P14>

Query Match 57.8%; Score 37; DB 1; Length 442;
Best Local Similarity 63.6%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 POGIACORNF 12
||||:|:|
DB 374 POGKAGQKGVN 384

RESULT 18

S47772
biotin sulfonamide reductase (EC 1.-.-.-) 1 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 01-Mar-2002
C:Accession: S47772; JV0071; A65154
R:Plunkett, G.
submitted to the EMBL Data Library, March 1994
A:Reference number: S47772
A:Accession: S47772
A:Molecule type: DNA
A:Residues: 1-739 <PIU>
A:Cross-references: EMBL:U00039; NID:9466582; PIDN:AA18528.1; PID:9466689
R:Pierston, D.E.; Campbell, A.
J. Bacteriol. 172, 2194-2198, 1990
A:Title: Cloning and nucleotide sequence of bscC, the structural gene for biotin sulf
A:Reference number: JV0071; MUID:90202748; PMID:2180932
A:Accession: JV0071
A:Molecule type: DNA
A:Residues: 1-544, 'ATPRSAGD', 554-557, 'OR', 561-708, 'MAVRYRRHGKNTTVR', <PIE>
A:Cross-references: GB:M44827; NID:g145435; PIDN:AA23522.1; PID:g145436
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.D.; Mau, B.; Zhao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: A65154
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-739 <BLAT>

A:Cross-references: GB:AE000432; GB:U00096; NID:g2367241; PIDN:AACT6575.1; PID:g1789973;
 A:Experimental source: strain K12, substrain MG1655
 C:Comment: This enzyme may serve as a scavenger, allowing the cell to utilize biotin sul
 C:Genetics:
 A:Gene: bsc
 A:Map position: 79 min
 C:Superfamily: trimethylamine-N-oxide reductase
 C:Keywords: ATP; molybdenum; nucleotide binding; oxidoreductase; P-loop
 F:486-493/Region: nucleotide-binding motif A (P-loop)

Query Match 57.8%; Score 37; DB 2; Length 739;
 Best Local Similarity 63.6%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPOGIAGORNF 11
 :|||||
 DB 34 NPOGIRGODEF 44

RESULT 19
 biotin sulfoxide reductase [imported] - Escherichia coli (strain O157:H7, substrain RMD
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: D91183
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gaswara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: D91183
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-739 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA37859.1; PID:g13363910; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0309952
 C:Genetics:
 A:Gene: ECS4436
 C:Superfamily: trimethylamine-N-oxide reductase

Query Match 57.8%; Score 37; DB 2; Length 739;
 Best Local Similarity 63.6%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPOGIAGORNF 11
 :|||||
 DB 34 NPOGIRGODEF 44

RESULT 20
 H86029
 biotin sulfoxide reductase [imported] - Escherichia coli (strain O157:H7, substrain EDL9
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H86029
 R:Perma, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Polomousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: H86029
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-739 <STO>
 A:Cross-references: GB:AE005174; NID:g12518286; PIDN:AAG58700.1; GSPDB:GN00145; UWGP:249
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: bsc
 C:Superfamily: trimethylamine-N-oxide reductase

Query Match 57.8%; Score 37; DB 2; Length 739;
 Best Local Similarity 63.6%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPOGIAGORNF 11
 :|||||
 DB 34 NPOGIRGODEF 44

RESULT 21
 564741
 cuticle collagen - tube worm (Riftia pachyptila) (fragments)

C:Species: Riftia pachyptila
 C:Date: 15-Nov-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
 C:Accession: S64741; S72264
 R:Mann, K.; Mechling, D.E.; Baechinger, H.P.; Eckerskorn, C.; Gall, F.; Timpl, R.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64741

A:Accession: S64741
 A:Molecule type: protein
 A:Residues: 1-42;43-86;89-148;149-206;207-265;266-320;321-403;404-446;447-489;490-542
 R:Mann, K.; Mechling, D.E.; Baechinger, H.P.; Eckerskorn, C.; Gall, F.; Timpl, R.
 J. Mol. Biol. 261, 255-266, 1996
 A:Title: Glycosylated threonine but not 4-hydroxyproline dominates the triple helix s
 A:Reference number: S72264; MUID:96346177; PMID:8757292

A:Accession: S72264
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-42;43-86;89-148;149-206;207-265;266-320;321-403;404-446;447-489;490-542
 C:Superfamily: unassigned collagens
 C:Keywords: extracellular protein

Query Match 57.8%; Score 37; DB 2; Length 751;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAQ 8
 :|||||
 DB 99 POGIAQ 105

RESULT 22
 E98161
 probable ATP-dependent DNA ligase PA2138 [imported] - Agrobacterium tumefaciens (stra
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: E98161
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quicillo, B.; Goldm
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz,
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
 A:Reference number: A97359; PMID:11743194
 A:Accession: E98161
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-771 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK8815.1; PID:g15158571; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L-502
 A:Map position: linear chromosome

Query Match 57.8%; Score 37; DB 2; Length 771;
 Best Local Similarity 70.0%; Pred. No. 66;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 POGIAQORNF 11
 :|||||
 DB 594 POGIAQORNF 603

RESULT 23
 AD3126
 ATP-dependent DNA ligase Atu4632 [imported] - Agrobacterium tumefaciens (strain C58,
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AD3126

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, H. erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McCell
A:Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193

A:Accession: AD3126

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-771 <KUR>

A:Cross-references: GB:AE008689; PIDN:AA145426.1; PID:g17743127; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

A:Genetics:

A:Gene: Atua4632

A:Map position: linear chromosome

Query Match

Best Local Similarity 57.8%; Score 37; DB 2; Length 771;
Pred. No. 66;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 POGIAGORNF 11

Db 594 POGIAGORNF 603

RESULT 24

B4033

C:Species: *Xenopus laevis* (African clawed frog)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: B40333

R:Su, M.W.; Suzuki, H.R.; Bieker, J.J.; Solursh, M.; Ramirez, F.
J. Cell Biol. 115, 565-575, 1991

A:Title: Expression of two nonallelic type II procollagen genes during *Xenopus laevis* em
A:Reference number: A40333; MUID:92011898; PMID:1918153

A:Accession: B40333

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1486 <SDA>

A:Cross-references: GB:M63595

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F:37-86/Domain: von Willebrand factor type C repeat homology <WMC>

F:1258-1486/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match

Best Local Similarity 57.8%; Score 37; DB 1; Length 1486;
Pred. No. 13e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9

Db 975 POGIAGOR 982

RESULT 25

T24797

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T24797

R:McMurry, A.
submitted to the EMBL Data Library, March 1997

A:Reference number: Z19936

A:Accession: T24797

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-369 <WIL>

A:Cross-references: EMBL:Z93388; PIDN:CA07663.1; GSPDB:GN00023; CESP:T10C6.7

A:Experimental source: clone T10C6

A:Map position: 5

A:Insertions: 44/2; 114/2; 219/3; 313/2

C:Superfamily: *Caenorhabditis elegans* hypothetical protein F08P3.6

Query Match

Best Local Similarity 57.0%; Score 36.5; DB 2; Length 369;
Pred. No. 38;

Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 1 SPQIAGORNF 12

Db 331 SPQIAGORNF-FN 341

RESULT 26

G83560

C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: G83560

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043

A:Accession: G83560

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-149 <STG>

A:Cross-references: GB:AE004503; GB:AE004091; MUID:g9946553; PIDN:AA004070.1; GSPDB:GN

A:Experimental source: strain PA01

A:Genetics:

A:Gene: PA0681

C:Superfamily: secretion protein xcpT

Query Match

Best Local Similarity 56.2%; Score 36; DB 2; Length 149;
Pred. No. 19;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPQIAGORNF 11

Db 9 SPQIAGORNF 19

RESULT 27

S16965

C:Species: *Yersinia pestis*

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S16965

R:Galayov, E.E.; Karlish, A.V.; Chernovskaya, T.V.; Dolgikh, D.A.; Smirnov, O.Y.; Vo
PERS Lett. 286, 79-82, 1991

A:Title: Expression of the envelope antigen F1 of *Yersinia pestis* is mediated by the
A:Reference number: S16965; MUID:91323540; PMID:1677900

A:Accession: S16965

A:Molecule type: DNA

A:Residues: 1-258 <GAL>

A:Cross-references: EMBL:X61966; MUID:948620; PIDN:CAA43967.1; PID:g1072423

A:Genetics:

A:Gene: cat1M

C:Superfamily: chaperone protein pnp

C:Keywords: molecular chaperone

F:1-20/Domain: signal sequence

F:21-258/Product: probable chaperone #status predicted <SIG>
F:121-160/Disulfide bonds: #status predicted <MAT>

Query Match

Best Local Similarity 56.2%; Score 36; DB 1; Length 258;
Pred. No. 33;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 POGIAGORNF 10

Db 228 POGIAGORNF 236

RESULT 33
AC2790
hypothetical protein Atu1738 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 17-May-2002
C:Accession: AC2790
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AC2790
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-577 <KUR>
A:Cross-references: GB:AE006688; PIDN:AL42737.1; PID:917740177; GSPDB:GN00186
C:Genetics:
A:Gene: Atu1738
A:Map position: circular chromosome
C:Superfamily: Escherichia coli ABC transporter mdla; Atp-binding cassette homology

Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 577;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PGCIAGORNF 11
Db 298 PKGIAGFRFR 307

RESULT 34
B97569
ABC transporter (ATP-binding protein) (AP001520) [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 17-May-2002
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quicillo, B.; Goldman,
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97559; PMID:11743194
A:Accession: B97569
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-598 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87507.1; PID:915156836; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3190
A:Map position: circular chromosome
C:Superfamily: Escherichia coli ABC transporter mdla; Atp-binding cassette homology

Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 598;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PGCIAGORNF 11
Db 319 PKGIAGFRFR 328

RESULT 35
C71905
probable outer membrane protein - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: C71905

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.
; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric F
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: C71905
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-657 <ARN>
A:Cross-references: GB:AE001497; GB:AE001439; MID:94155199; PIDN:AND06225.1; PID:9415
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0649

Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 657;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPGIAGORNF 12
Db 558 SPGIOTKRNF 569

RESULT 36
F64608
conserved hypothetical protein HP0710 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: F64608
R:Tom, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glöck, A.; McK
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weisman, J.M.; Fujii, C.; Bowman, C.; Wetthey,
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: F64608
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-660 <TON>
A:Cross-references: GB:AE000584; GB:AE000511; MID:92313834; PIDN:AND0765.1; PID:9231

Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 660;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPGIAGORNF 12
Db 561 SPGIOTKRNF 572

RESULT 37
E71879
probable outer membrane protein - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.
; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric F
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: E71879
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-668 <ARN>
A:Cross-references: GB:AE001516; GB:AE001439; MID:94155431; PIDN:AND06437.1; PID:9415
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0870

Query Match
Best Local Similarity 56.2%; Score 36; DB 2; Length 668;

Best Local Similarity 58.3%; Pred. No. 89;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPQGIAGQNFEN 12
||| :|||:
DB 569 SPTGIQTKNFES 580

RESULT 38

AD0982
biotin sulfoxide reductase (EC 1.-.-.-) [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AD0982
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Mature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AD0982
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-777 <PAR>
A:Cross-References: GB:AL513382; PIDN:CAD07984.1; PID:q16504970; GSPDB:GN00176
C:Genetics:
A:Gene: bsc
C:Superfamily: trimethylamine-N-oxide reductase
C:Keywords: oxidoreductase

Query Match 56.2%; Score 36; DB 2; Length 777;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 POGIAGQNF 11
||| | | |
DB 73 POGIAGQNF 82

RESULT 39

T06494
1,4-alpha-glucan branching enzyme (EC 2.4.1.18) II - garden pea (fragment)
N:Alternate names: starch branching enzyme II
C:Species: Pisum sativum (garden pea)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T06494
R:Burton, R.A.; Bewley, J.D.; Smith, A.M.; Bhattacharya, M.K.; Tatge, H.; Ring, S.; Bul
Plant J. 7, 3-15, 1995
A:Title: Starch branching enzymes belonging to distinct enzyme families are differentia
A:Reference number: Z15717; MUID:95201826; PMID:7894509
A:Accession: T06494
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-826 <BUR>
A:Cross-References: EMBL:X80010; NID:9510546; PIDN:CAA56320.1; PID:q1345571
C:Genetics:
A:Note: SBEII
C:Superfamily: 1,4-alpha-glucan branching enzyme
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 56.2%; Score 36; DB 2; Length 826;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 1 SPQGIAG--QNFEN 12
||| | | : | | |
DB 725 SPQGIAGIPIETNFEN 728

RESULT 40

T07824

1,4-alpha-glucan branching enzyme (EC 2.4.1.18) I (clone SBE17) - potato (fragment)
C:Species: Solanum tuberosum (potato)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07824
R:Khoshnoodi, J.; Blenow, A.; Ek, B.; Rask, L.; Larsson, H.
Eur. J. Biochem. 242, 148-155, 1996
A:Title: The multiple forms of starch branching enzyme I in Solanum tuberosum.
A:Reference number: Z16155; MUID:97112484; PMID:8954164
A:Accession: T07824
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-830 <KHO>
A:Cross-References: EMBL:X08786; NID:q1621011; PIDN:CAV0039.1; PID:q1621012
A:Experimental source: cv. Diamella; cell line Diamella
C:Genetics:
A:Gene: SBEI
C:Function:
A:Description: catalyzes the transglycosylation of a terminal segment of a 1,4-alpha-

A:Pathway: glycogen/starch biosynthesis
C:Superfamily: 1,4-alpha-glucan branching enzyme
C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 56.2%; Score 36; DB 2; Length 830;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 1 SPQGIAG--QNFEN 12
||| | | : | | |
DB 669 SPQGIAGIPIETNFEN 682

RESULT 41

S34730
1,4-alpha-glucan branching enzyme (EC 2.4.1.18) precursor, amyloplast - potato
N:Alternate names: starch branching enzyme
C:Species: Solanum tuberosum (potato)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 07-Dec-1999
C:Accession: S34730; S38733; S18594
R:Poulsen, P.
submitted to the EMBL Data Library, December 1992
A:Description: Starch-branching enzyme cDNA from Solanum tuberosum.
A:Reference number: S34730
A:Accession: S34730
A:Molecule type: mRNA
A:Residues: 1-861 <POU>
A:Cross-References: EMBL:X68805; NID:9396080; PIDN:CAA9463.1; PID:q396081
R:Khoshnoodi, J.; Ek, B.; Rask, L.; Larsson, H.
FEBS Lett. 332, 132-138, 1993
A:Title: Characterization of the 97 and 103 kDa forms of starch branching enzyme from
A:Reference number: S38732; MUID:94009663; PMID:8405428
A:Accession: S38733
A:Status: preliminary
A:Molecule type: protein
A:Residues: 76-95;236-244, 'X', 246-255;311-329;393-402;515-520;523-529;545-558;'F', 636
A:Accession: S38732
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 318-492;'S', 494-538, 'K', 540-550 <KHO>
R:Kossmann, J.; Visser, R.G.F.; Mueller-Roeber, B.; Willmitzer, L.; Sonnwald, U.
Mol. Gen. Genet. 230, 39-44, 1991
A:Title: Cloning and expression analysis of a potato cDNA that encodes branching enzy
A:Reference number: S18594; MUID:92079917; PMID:1745241
A:Accession: S18594
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 279-527 <KOS>
C:Genetics:
A:Gene: SBE
A:Genome: nuclear
C:Function:
A:Description: catalyzes the transglycosylation of a terminal segment of a 1,4-alpha-

A:Pathway: glycogen/starch biosynthesis
A:Note: final step in biosynthesis of glycogen or amylopectin

C:Superfamily: 1,4-alpha-glucan branching enzyme
 C:Keywords: amyloplast; glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase
 F:1-75/Domain: transit peptide (amyloplast) #status predicted <TMP>
 F:76-861/Product: 1,4-alpha-glucan branching enzyme #status experimental <MAT>

Query Match
 Best Local Similarity 56.2%; Score 36; DB 1; Length 861;
 Matches 8; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

OY 1 SPOGIAG--ORNFN 12
 ||:| | : ||
 Db 744 SPEGIRGVETNFN 757

RESULT 42

androgen receptor A - human
 C:Species: Homo sapiens (man)
 C>Date: 31-Mar-1991 #sequence-revision 31-Mar-1991 #text-change 12-Sep-1997
 C:Accession: A34721

Mol. Endocrinol. 4, 417-427, 1990
 A:Title: Specific region in hormone binding domain is essential for hormone binding and
 A:Reference number: A34721; MUID:90258935; PMID:2342476
 A:Accession: A34721

A:Molecule type: mRNA
 A:Residues: 1-910 <GOV>
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: zinc finger
 F:548-806/Domain: erba transforming protein homology <ERBA>
 F:550-570/Region: zinc finger
 F:586-610/Region: zinc finger

Query Match
 Best Local Similarity 56.2%; Score 36; DB 2; Length 910;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 POGIAGORN 10
 ||:| | : ||
 Db 477 POGIAGORN 485

RESULT 43

androgen receptor B - human
 C:Species: Homo sapiens (man)
 C>Date: 31-Mar-1991 #sequence-revision 31-Mar-1991 #text-change 12-Sep-1997
 C:Accession: B34721

Mol. Endocrinol. 4, 417-427, 1990
 A:Title: Specific region in hormone binding domain is essential for hormone binding and
 A:Reference number: A34721; MUID:90258935; PMID:2342476
 A:Accession: B34721

A:Molecule type: mRNA
 A:Residues: 1-911 <GO2>
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: zinc finger
 F:548-806/Domain: erba transforming protein homology <ERBA>
 F:550-570/Region: zinc finger
 F:586-610/Region: zinc finger

Query Match
 Best Local Similarity 56.2%; Score 36; DB 2; Length 911;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 POGIAGORN 10
 ||:| | : ||
 Db 477 POGIAGORN 485

RESULT 44

A39248
 androgen receptor - human

C:Species: Homo sapiens (man)
 C>Date: 04-Oct-1991 #sequence-revision 04-Oct-1991 #text-change 24-Nov-1999
 C:Accession: A39248; A30328; A40109; A60946; A34942; A27653; A40108; A40494; A32224;
 R:Rubin, D.B.; Brown, T.R.; Stamenkovic, J.A.; Higgins, H.N.; Migeon, C.J.; Wilson, E.M.;
 Proc. Natl. Acad. Sci. U.S.A. 86, 9534-9538, 1989
 A:Title: Sequence of the intron/exon junctions of the coding region of the human andr
 A:Reference number: A39248; MUID:90083302; PMID:2594783
 A:Accession: A39248

A:Molecule type: DNA
 A:Residues: 1-919 <LUB>
 A:Cross-references: GB:M27423; GB:M27430; NID:9178904; PID:AAA51886.1; PID:9178906
 R:Faber, P.W.; Kuiper, G.G.J.M.; van Rooij, H.C.J.; van der Korp, J.A.G.M.; Brinkma
 Mol. Cell. Endocrinol. 61, 257-262, 1989

A:Title: The N-terminal domain of the human androgen receptor is encoded by one, larg
 A:Reference number: A30328; MUID:89137730; PMID:2917688
 A:Accession: A30328

A:Molecule type: DNA
 A:Residues: 1-77, 79-165, 'A', 167-389, 'L', 391-464, 473-538 <FAB>
 A:Cross-references: GB:M20260
 R:Rubin, D.B.; Joseph, D.R.; Sullivan, P.M.; Willard, H.F.; French, F.S.; Wilson, E.

Science 240, 327-330, 1988
 A:Title: Cloning of human androgen receptor complementary DNA and localization to the
 A:Reference number: A40109; MUID:88178112; PMID:3353727
 A:Accession: A40109

A:Molecule type: DNA
 A:Residues: 559-624 <LUB>
 A:Cross-references: GB:M20132
 R:Kuiper, G.G.J.M.; Faber, P.W.; van Rooij, H.C.J.; van der Korp, J.A.G.M.; Rits-Ste
 J. Mol. Endocrinol. 2, R1-R4, 1989

A:Title: Structural organization of the human androgen receptor gene.
 A:Reference number: A60946; MUID:89322749; PMID:2546571
 A:Accession: A60946

A:Molecule type: DNA
 A:Residues: 536-540; 587-591; 626-631; 723-726; 770-774; 814-818; 867-870 <KUI>
 R:Rubin, D.B.; Joseph, D.R.; Say, M.; Tan, J.; Higgins, H.N.; Larson, R.E.; French, F.
 Mol. Endocrinol. 2, 1265-1275, 1988

A:Title: The human androgen receptor: complementary deoxyribonucleic acid cloning, se
 A:Reference number: A34942; MUID:89112208; PMID:3216866
 A:Accession: A34942

A:Molecule type: mRNA
 A:Residues: 1-919 <LUB>
 A:Cross-references: GB:M20132; NID:9178627; PID:AAA51729.1; PID:9178628; GB:J03180
 R:Trapani, J.; Klaassen, P.; Kuiper, G.G.J.M.; van der Korp, J.A.G.M.; Faber, P.W.;
 Biochem. Biophys. Res. Commun. 153, 241-248, 1988

A:Title: Cloning, structure and expression of a cDNA encoding the human androgen rece
 A:Reference number: A27653; MUID:88240407; PMID:3377788
 A:Accession: A27653

A:Molecule type: mRNA
 A:Residues: 468-564, 'K', 566-919 <TRA>
 A:Cross-references: GB:M20260; NID:9178891; PID:AAA51774.1; PID:9178892
 A:Note: the authors translated the codon AAG for residue 565 as Glu
 R:Chang, C.; Kokontis, J.; Liao, S.
 Science 240, 324-326, 1988

A:Title: Molecular cloning of human and rat complementary DNA encoding androgen recep
 A:Reference number: A40108; MUID:88178111; PMID:3353726
 A:Accession: A40108

A:Molecule type: mRNA
 A:Residues: 557-628 <CHA>
 A:Cross-references: GB:M18624
 R:Chang, C.; Kokontis, J.; Liao, S.
 Proc. Natl. Acad. Sci. U.S.A. 85, 7211-7215, 1988

A:Title: Structural analysis of complementary DNA and amino acid sequences of human a
 A:Reference number: A40494; MUID:89017168; PMID:3174628
 A:Accession: A40494

A:Molecule type: mRNA
 A:Residues: 1-74, 79-89, 'H', 90-472, 'GGG', 473-474, 'E', 476-644, 'N', 646-919 <CH2>
 A:Cross-references: GB:M23263
 R:Trilley, W.D.; Marcelletti, M.; Wilson, J.D.; McPhaul, M.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 327-331, 1989
 A:Title: Characterization and expression of a cDNA encoding the human androgen recept
 A:Reference number: A32224; MUID:99098909; PMID:2911578
 A:Accession: A32224

```

Query Match          56.2%; Score 36; DB 2; Length 919;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      2 POGINGORN 10
         |||:||||:
db       486 POGIAGOS 494

```

Query Match	56.2%;	Score 36;	DB 2;	Length 1354;
Best Local Similarity	100.0%;	Pred. No. 1.9e+02;		
Matches	7;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

RESULT 46

collagen alpha 1(IV) chain precursor - fruit fly (*Drosophila melanogaster*)
A1595
C:Species: *Drosophila melanogaster*
C:Date: 21-May-1990 #sequence_id:1990 #text_change 21-Jul-2000
C:Accession: A31893; A26992; A19442; S00020
R:Blumberg, B.; McKrell, A.J.; Feeler, J.H.
J. Biol. Chem. 265, 18328-18337, 1988
Title: *Drosophila* basement membrane procollagen alpha-1(IV). II. Complete cDNA sequence

Query Match	56.2%;	Score 36;	DB 2;	Length 1775;
Best Local Similarity	54.5%;	Pred. No. 2.5e+02;		
Matches	6;	Conservative	1;	Mismatches 4; Indels 0; Gaps 0;

RESULT 47

12/7367
 hypothetical protein Y73FA.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27367
 R:Steward, C.
 submitted to the EMBL data Library, October 1998
 A:Reference number: Z20358
 A:Accession: T27367
 A:Status: preliminary; translated from GB/EMBL/DDAJ
 A:Molecule type: DNA

A:Residues: 1-182 <WIL>
 A:Cross-references: EMBL:AL032661; PIDN:CAA21755.1; CESP:V73F4A.1
 A:Experimental source: clone Y73F4A
 C:Genetics:
 A:Gene: CESP:V73F4A.1
 A:Introns: 64/3; 161/3

Query Match
 Best Local Similarity 54.7%; Score 35; DB 2; Length 182;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 OY 2 PGAGGRRNF 12
 DB 120 PGAGGRRNF 130

RESULT 48
 T24078
 hypothetical protein R09D1.7 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T24078
 R:Matthews, P.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: 219838
 A:Accession: T24078
 A:Status: preliminary; translated from GB/EMBL/DDB
 A:Molecule type: DNA
 A:Residues: 1-369 <WIL>
 A:Cross-references: EMBL:Z70035; PIDN:CAA93867.1; GSPDB:GN00020; CESP:R09D1.7
 A:Experimental source: clone R09D1
 C:Genetics:
 A:Gene: CESP:R09D1.7
 A:Map position: 2
 A:Introns: 5/1; 97/2

Query Match
 Best Local Similarity 54.7%; Score 35; DB 2; Length 369;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 4 GIAGRRNF 12
 DB 210 GVIGRRNF 218

RESULT 49
 H96933
 aspartate kinase [imported] - *Clostridium acetobutylicum*
 C:Species: *Clostridium acetobutylicum*
 C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: H96933
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, O.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: H96933
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-437 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK78259.1; PID:915023117; GSPDB:GN00168
 A:Experimental source: *Clostridium acetobutylicum* ATCC824
 C:Genetics:
 A:Gene: CAC0278

Query Match
 Best Local Similarity 54.7%; Score 35; DB 2; Length 437;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 4 GIAGRRNF 11
 DB 290 GIAGRRNF 297

RESULT 50
 AB0129
 Probable membrane protein YPO1051 [imported] - *Yersinia pestis* (strain CO92)
 C:Species: *Yersinia pestis*
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
 C:Accession: AB0129
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchall, R.W.; Holden, M.T.G.; Prentice, M.; deno-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AB0129
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-451 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC89893.1; PID:915979118; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO1051
 C:Superfamily: Escherichia coli probable zinc proteinase yael

Query Match
 Best Local Similarity 54.7%; Score 35; DB 2; Length 451;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 OY 1 SPQGIAGRRNF 12
 DB 135 SPQSIAGRRNF 146

Search completed: May 16, 2003, 10:41:19
 Job time : 29 secs

